Available online at www.sciencedirect.com
SCIENCE DIRECT®

Carbohydrate Research 339 (2004) 937–948

Carbohydrate
RESEARCH

Evaluation of carbohydrate molecular mechanical force fields by quantum mechanical calculations

Lars Hemmingsen,[a,*] Daniel E. Madsen,[b] Anders L. Esbensen,[b] Lars Olsen[b] and Søren B. Engelsen[c]

[a]*Quantum Protein Centre, Department of Physics, The Technical University of Denmark, Bldg 309, DK-2800 Lyngby, Denmark*
[b]*Department of Mathematics and Physics, The Royal Veterinary and Agricultural University, Thorvaldsensvej 40, DK-1871, Denmark*
[c]*Food Technology, The Royal Veterinary and Agricultural University, Rolighedsvej 30, DK-1958 Frederiksberg C, Denmark*

Received 8 August 2003; accepted 20 November 2003

Abstract—A quantitative evaluation of 20 second-generation carbohydrate force fields was carried out using ab initio and density functional methods. Geometry-optimized structures (B3LYP/6-31G(d)) and relative energies using augmented correlation consistent basis sets were calculated in gas phase for monosaccharide carbohydrate benchmark systems. Selected results are: (i) The interaction energy of the α-D-glucopyranose·H_2O heterodimer is estimated to be 4.9 kcal/mol, using a composite method including terms at highly correlated (CCSD(T)) level. Most molecular mechanics force fields are in error in this respect; (ii) The $_3E$ envelope (*south*) pseudorotational conformer of methyl 5-deoxy-β-D-xylofuranoside is 0.66 kcal/mol more stable than the ^3E envelope (*north*) conformer and the α-anomer of methyl D-glucopyranoside is 0.82 kcal/mol more stable than the β-anomer; (iii) The relative energies of the (*gg*, *gt* and *tg*) rotamers of methyl α-D-glucopyranoside and methyl α-D-galactopyranoside are (0.13, 0.00, 0.15) and (0.64, 0.00, 0.77) kcal/mol, respectively. The results of the quantum mechanical calculations are compared with the results of calculations using the 20 second-generation carbohydrate force fields. No single force field is consistently better than the others for all the test cases. A statistical assessment of the performance of the force fields indicates that CHEAT(95), CFF, certain versions of Amber and of MM3 have the best overall performance, for these gas phase monosaccharide systems.

Keywords: DFT; Ab initio; Force field; Methyl 5-deoxy-β-D-xylofuranoside; Methyl β-D-glucopyranoside; Methyl α-D-glucopyranoside; Methyl α-D-galactopyranoside; α-D-Glucopyranose monohydrate; Hydration energy

1. Introduction

Carbohydrates constitute a group of molecules with diverse biological functions and molecular modeling of carbohydrates, or carbohydrate-containing molecules, has become increasingly important for pharmaceutical and material science applications, considering the prominent role of carbohydrates in, for example, human immunochemistry and cell–cell recognition/adhesion. However, as flexible polyhydroxy compounds, carbohydrates display a multiconformational behavior in the predominant biological medium water, which seriously complicates molecular modeling. The high structural flexibility of carbohydrates and their ability to ring open and close in solution creates 'room' for a significant interplay with subtle electronic effects that occur during conformational and configurational changes, such as the anomeric, exo-anomeric, and gauche effects[1] and directional hydrogen bonding interactions with semi-structural water molecules.[2]

Due to these inherent complex carbohydrate properties, the development of adequate carbohydrate force fields is a nontrivial task. The original general purpose force fields such as MM2,[3] CHARMM,[4] AMBER,[5] and GROMOS[6] developed with polypeptides and nucleic acids in mind all displayed poor performance when

---

* Corresponding author. Tel.: +45-45253281; fax: +45-45931669; e-mail: lhemming@fysik.dtu.dk

0008-6215/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2003.11.024

applied to carbohydrates, for which reason there has been a tradition of re-parameterizations and parameterizations of new carbohydrate force fields that would account for the specific features of carbohydrates. In a recent investigation by Pérez et al.[7] the performance of 20 second-generation molecular mechanics carbohydrate force fields was compared on a relative basis. The benchmark was performed on a series of test molecules in the gas phase as simple one-conformer calculations with and without geometry optimization. The results cannot be unambiguously compared to ensemble-averaged experimental data obtained in condensed phase, for which reason it was decided to compare the force field performances on a relative basis using principal component and cluster analysis. This comparative analysis revealed a relatively wide spectrum of answers to the posed modeling tasks and the solutions provided could classify the individual carbohydrate force fields into one of several force field groups/families. One of the conclusions made was that the end users should always aim at investigating the problem in at least two different force fields from different families.

As the molecular problems investigated in the relative force field benchmark (as most other, if not all, single structure calculations of carbohydrates) have not been studied experimentally in the gas phase, the only feasible way to further validate the carbohydrate force fields is to perform quantum mechanical calculations at the highest possible level, which is the aim of the present study. During the past decade, it has been demonstrated that ab initio and density functional (DFT) methods can produce results, for small molecules in the gas phase, which are sufficiently reliable to be used in tests and parameterizations of force fields.[8–37] In this work, we set out to provide an absolute scale for the molecular properties from the test set, and thereby an absolute criterion for evaluating the performance of the force fields tested by Pérez et al.[7] The model systems in this benchmark comprise both mono- and disaccharides, but only the monosaccharide systems are included in this work (see Fig. 1). This study will focus on three apparent problems identified in the force field benchmark test that ironically are most important generic carbohydrate traits: (i) Anomeric ratio. Most force fields predicted a small energy preference for the β-anomer in the investigated two-state system; (ii) Hydroxyl rotamer preferences: Most force fields generally have reversed the ranking observed in solution *gg* > *gt* > *tg* in the investigated three-state system for methyl α-D-glucopyranoside and: (iii) Interaction energy of the α-D-glucopyranose·H_2O heterodimer: This energy varies from 0.85 to 13.72 kcal/mol using the 20 force fields.

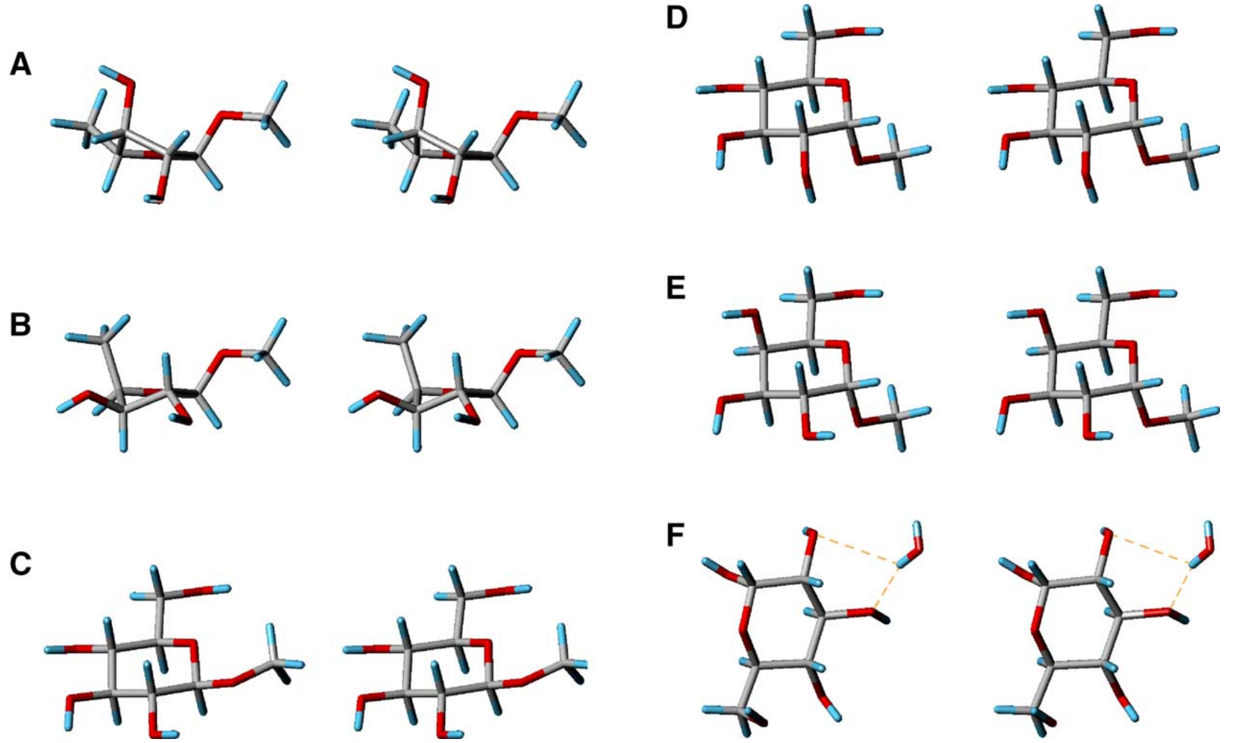

Figure 1. 3D Stereo drawing of the carbohydrate molecules used in the present study. (A) Methyl 5-deoxy-β-D-xylofuranoside (*south*), (B) methyl 5-deoxy-β-D-xylofuranoside (*north*), (C) methyl β-D-glucopyranoside, (D) methyl α-D-glucopyranoside, (E) methyl α-D-galactopyranoside, and (F) α-D-glucopyranose·H_2O.

2. Methods

All the calculations were carried out using the GAUSSIAN 98 program[38] on PCs. We carried out single point energy calculations with convergence criteria of 10^{-8} and 10^{-6} au for the maximal and rms change of the density matrix elements, respectively, and geometry optimizations with the default convergence criteria of GAUSSIAN 98. The basis sets used were 6-31G(d), 6-311+G(d,p), cc-pVXZ, and AUG-cc-pVXZ (X = D,T,Q,5)[39–49] and the methods HF, MP2, CCSD(T), and B3LYP.[50–54] If not otherwise stated, the B3LYP/AUG-cc-pVDZ//B3LYP/6-31G(d) method was applied. Initial guesses for carbohydrate structures were constructed mildly pre-optimized as described by Pérez et al.[7] Only in the case of the α-D-glucopyranose·H_2O system was the crystal structure[55] used directly, and the structure used by Halkier et al.[56] was adopted for the water dimer. In the H_2O·H_2O system and the α-D-glucopyranose·H_2O system the interaction energy was calculated, taking into account the basis set superposition error (BSSE) by performing a counterpoise correction.[57] The interaction energies and relative energies presented in this work are derived directly from the single point energy calculations, that is, no zero point energy or other corrections are added.

Principal component analysis (PCA)[58] was applied for gaining an overview of systematic variations in the force field performances. A two-dimensional data evaluation strategy was used in which many variables are registered (molecular geometry and energetics, in the present case) for the same objects (force fields, in the present study). PCA gives a representation, in a graphic interface, of the main variance of the complex data structure by projecting the data onto a few orthogonal directions (principal components). The scores contain systematic information about the objects and the loadings contain systematic information about the variables. In order to project the entire force field variation into one two-dimensional plot, a Sammon map[59] was constructed from the PCA score vectors. The Sammon map is a relatively simple nonlinear mapping procedure in which the (low) two-dimensional distances are optimized to mimic the N-dimensional distances according to the equation:

$$\min E = \frac{1}{\sum_{i<j}^{N} D_{ij}} \sum_{i<j}^{N} \frac{(d_{ij} - D_{ij})^2}{D_{ij}}$$

where d_{ij} are the distances in the lower dimensional space and the D_{ij} are the distances in the original higher dimensional space. N is the number of objects.

The multivariate data analysis was performed using the chemometric program Unscrambler 7.6 (CAMO, Trondheim, Norway) and Matlab 6.1 (The Mathworks Inc., Natick, MA, USA).

3. Results and discussions

3.1. Method and accuracy

Carbohydrate molecules, even monosaccharides, are too large to allow for highly correlated methods to be applied together with large basis sets. The B3LYP density functional method currently seems to be the optimal choice for relatively large systems, and therefore an evaluation of the quality of this method for the monosaccharide systems studied in this work was performed. Since intramolecular hydrogen bonding is an important factor for these systems, a suitable choice of a very small model system is the water dimer. In particular, this is relevant for the α-D-glucopyranose·H_2O system. A similar approach was used in the work of Lii et al.[12] Halkier et al.[56] performed systematic calculations of the interaction energy of the water dimer using ab initio methods (HF to CCSD(T)) and large basis sets (cc-pVDZ to dAUG-cc-pVQZ), providing benchmark data with which to compare our results. A systematic test of the effect of using different size basis sets on the interaction energy was performed using the B3LYP method (see Table 1). The most striking feature is that inclusion of diffuse functions in the basis set is extremely cost effective for the calculation of ΔE; for example, the AUG-cc-pVDZ basis set gives a result, which is slightly better than cc-pV5Z (as compared to the best basis set) at a computational time, which is about 200 times shorter. Csonka[60] similarly recommends the inclusion of diffuse functions, at least on the heavy atoms. Furthermore, the basis set superposition error (BSSE) is only 0.2 kcal/mol using this basis set, making it the optimal choice for the calculations in this work. Moreover (Table 1), the counterpoise correction is observed to be reliable in this case, even for the smallest basis set, making all the corrected interaction energies very similar

Table 1. Interaction energy of the water dimer

Basis set	ΔE (kcal/mol)	ΔE^{CP} (kcal/mol)	Relative CPU time
cc-pVDZ	−8.00	−4.46	1
cc-pVTZ	−5.97	−4.44	8
cc-pVQZ	−5.23	−4.45	90
cc-pV5Z	−4.69	−4.45	1100
AUG-cc-pVDZ	−4.59	−4.35	6
AUG-cc-pVTZ	−4.46	−4.41	43
AUG-cc-pVQZ	−4.47	−4.44	570
AUG-cc-pV5Z	−4.45	−4.45	5500

B3LYP calculations with different basis sets of the water dimer interaction energy (uncorrected and CP corrected). The geometry is fixed; see the Methods section.

(within 0.1 kcal/mol). It may be questioned if this test is also sufficient for the selection of basis set for the larger monosaccharides. We therefore performed an additional test of both the basis set and the B3LYP method. The interaction energy of the α-D-glucopyranose·H_2O system (see Tables 2 and 3). In Table 2, results of B3LYP calculations are presented with AUG-cc-pVDZ on the full system and either the AUG-cc-pVDZ or AUG-cc-pVTZ basis set for two models of the system (see Fig. 2). The CP correction is larger for these systems, 0.5–0.7 kcal/mol, than for the water dimer using the AUG-cc-pVDZ basis set. For the α-D-glucopyranose·H_2O system, the results in Table 2 show that the counterpoise corrected values with the AUG-cc-pVDZ basis set are within 0.1 kcal/mol of the corrected energies when using the larger AUG-cc-pVTZ basis set, indicating that the CP-corrected energy determined with the AUG-cc-pVDZ basis set is a good estimate of the complete basis set limit at the B3LYP level. Nevertheless, a larger error may apply when calculating relative energies of systems, which have a different number of hydrogen bonds. Surprisingly, truncating the molecule does not have a great effect on the calculated interaction energy, indicating that the most important interactions are represented in the smallest system. That is, local interactions (hydrogen bonds) dominate the interaction and thus dipole–dipole and other interactions between the water molecule and OH groups (not present in the smallest systems) of the glucose molecule only appear to give a minor contribution to the interaction energy. This allows for the use of the smallest system as a test case where higher level methods can be applied. In order to have a full comparison with ab initio methods, we performed HF, MP2, and CCSD(T) calculations on this system (see Table 3). The characteristics are basically the same for the model system of α-D-glucopyranose·H_2O as for the water dimer: HF underestimates the interaction energy, while MP2 and B3LYP improve the results, but still underestimate the interaction energy, getting within 0.9 kcal/mol of the CCSD(T) values. The convergence toward the complete basis set limit is slower for MP2 and CCSD(T) than for HF and B3LYP (see Table 1 and Halkier et al.[56]) and, accordingly, the CP corrections are larger for MP2 and CCSD(T). For the CCSD(T) method, the results of Halkier et al. even show that the CP method overcorrects the energy and that a better estimate of the complete basis set limit energy is the average of the uncorrected and the corrected energies.

Based on the discussion above, we estimate that the B3LYP/AUG-cc-pVDZ method is correct within 1 kcal/mol for the calculation of the energy of a hydrogen bond, when the energy is corrected for BSSE. It seems reasonable to assume that the method performs better than this for the relative energies of different conformers

Table 2. Interaction energy of different model systems of α-D-glucopyranose·H_2O heterodimer using the B3LYP method and different basis sets (uncorrected and CP-corrected) (see Fig. 2)

System/Basis set	ΔE (kcal/mol)	ΔE^{CP} (kcal/mol)
Full system		
AUG-cc-pVDZ	−3.85	−3.18
Model system 1		
AUG-cc-pVDZ	−4.13	−3.51
AUG-cc-pVTZ	−3.70	−3.62
Model system 2		
AUG-cc-pVDZ	−3.96	−3.41
AUG-cc-pVTZ	−3.58	−3.44

Table 3. Interaction energy of the water dimer and of model system 2 (see Fig. 2) of the α-D-glucopyranose·H_2O heterodimer, respectively, for different methods with the AUG-cc-pVDZ basis set (uncorrected and CP-corrected)

Method	Glc·H_2O		H_2O·H_2O	
	ΔE (kcal/mol)	ΔE^{CP} (kcal/mol)	ΔE (kcal/mol)	ΔE^{CP} (kcal/mol)
HF	−3.29	−2.71	−3.75[a]	−3.50[a]
B3LYP	−3.96	−3.41	−4.59	−4.35
MP2	−5.11	−3.56	−5.16[a]	−4.30[a]
CCSD(T)	−5.90	−4.27	−5.20[a]	−4.23[a]

[a] From Halkier et al.[56]

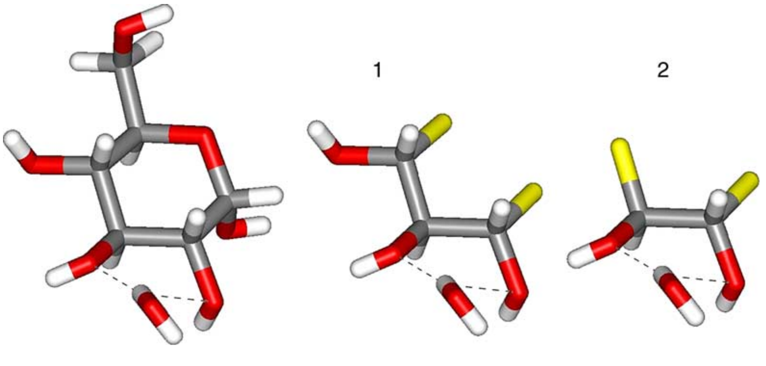

Figure 2. Different model systems of the α-D-glucopyranose·H_2O heterodimer. Full system (left), model system 1 (center), and model system 2 (right). Broken C–C bonds are saturated with hydrogen atoms shown in yellow.

of a given monosaccharide, if the number of hydrogen bonds in the different conformers is the same. This is supported by comparing the relative energy of the *gg*, *tg* and *gt* conformers of methyl α-D-glucopyranoside found in this work and the same relative energies found using a composite energy including correlation effects to high level in the work by Barrows et al.[11] The relative energies found here are within 0.2 kcal/mol of those found by Barrows et al. Note that the comparison is not fully consistent, as Barrows et al. studied glucose and not the methylated glucoside.

We have applied the B3LYP/6-31G(d) method for the geometry optimizations in this work. The same approach was used by Momany and Willett[8] who found it to give results very similar to B3LYP/6-31++G(d,p) and by Gordon et al.[28] and Tvaroska and Carver[14] who found essentially the same relative energies for structures optimized with B3LYP/6-31G(d) and B3LYP/6-31+G(d,p) for methyl β-D-arabinofuranoside and fluoro and chloro derivatives of tetrahydropyran, respectively. Almost the same approach was used by Lii et al.[12] who found that the relative energies of *gg*, *gt*, and *tg* conformers of glucopyranose change by less than 0.06 kcal/mol when applying B3LYP/6-31++G(2d,2p)//B3LYP/6-31++G(2d,2p) instead of B3LYP/6-31++G(2d,2p)//B3LYP/6-31G(d,p). Moreover, the results of Chung-Phillips and Chen[13] show that the method for optimization does not have a large effect on the calculated relative energy of the α- and β-anomers of 2-hydroxytetrahydrofuran. The energies calculated with MP2/6-31G(d,p)//HF/6-31G(d) and MP2/6-31G(d,p)//MP2/6-31G(d,p) differ by 0.04 kcal/mol and HF/6-311++G(d,p)//HF6-31G(d) and HF/6-311++G(d,p)//MP2/6-31+G(d,p) differ by 0.11 kcal/mol. In addition, we calculated the relative energy of *north* (the ^3E envelope) and *south* (the $_3E$ envelope) conformers of methyl 5-deoxy-β-D-xylofuranoside using two methods: B3LYP/AUG-cc-pVDZ//B3LYP/6-31G(d) and B3LYP/AUG-cc-pVDZ//B3LYP/6-311+G(d,p). The relative energy changed by only 0.12 kcal/mol, but some changes in the two Cremer–Pople pucker parameters[61] describing the conformation of the furanose ring are observed for the *north* conformer (see Table 4). Thus, the error originating from the imperfection of geometry optimization is apparently significantly smaller than what was estimated above for the calculation of the energy.

As a rough estimate of the total error on the calculated relative energies presented in this work, we assume a contribution smaller than 1 kcal/mol due to the imperfection of the energy calculation and a contribution of about 0.1 kcal/mol from imperfection of the geometry optimization. Thus, the total estimated error is smaller than 1 kcal/mol. However, this estimate should be validated in a thorough analysis of the accuracy of the relative energies of different conformers.

3.2. Evaluation of molecular mechanics force field benchmark

Based on the results outlined above, we applied B3LYP/AUG-cc-pVDZ for energy calculations and B3LYP/6-31G(d) for geometry optimizations to the test cases of the monosaccharide carbohydrate benchmark.[7] Figure 1 displays graphical representations of most of the molecules investigated here. The choice of model systems is related to essential carbohydrate features such as puckering magnitude and preference for furanose rings, anomer conformation and stability of α and β hexopyranose conformers, primary hydroxyl rotamer preferences for hexopyranoses with and without neighboring axial group and the interaction energy of α-D-glucopyranose·H_2O system. Some of the force fields, and in particular the Amber GB/SA and MM3*(80), are designed for condensed phase calculations. Therefore, results obtained using these force fields are not directly comparable to gas phase QM data. Basically, the dielectric constant is a measure of the modification of the Coulomb interactions due to the intervening medium. While the dielectric constant is well defined in vacuum or in a solvent like water, it is not straight forward to choose a meaningful value for intra-molecular interactions. Most force fields use a dielectric constant of 1 for both intra- and inter-molecular interactions, embodying the effective dielectric constant in the fractional charges. Some of the force fields in this work have relative dielectric constants different from 1, though most use values of less than 5. These values are not chosen to model solvent implicitly, but represent a convenient scaling of the Coulomb interactions, which allows for the best reproduction of various experimental data. Thus, we also include force fields with dielectric constants larger than 1 in this study.

Table 4. Cremer–Pople pucker parameters and relative energy of the *north* and *south* conformers of methyl 5-deoxy-β-D-xylofuranoside

Basis set	North (^3E)		South ($_3E$)		ΔE_{S-N} (kcal/mol)
	q_2	φ_2 (deg)	q_2	φ_2 (deg)	
6-31G(d)	0.34	−85.0	0.38	127.9	−0.79
6-311+G(d,p)	0.32	−71.1	0.38	128.4	−0.66

Geometry optimization was performed with the basis set in the first column and the B3LYP method, and energy evaluation at the B3LYP/AUG-cc-pVDZ level.

3.3. Interaction energy of the α-D-glucopyranose·H_2O heterodimer

Determination of the interaction energy between a simple monosaccharide, such as α-D-glucopyranose, and one water molecule may appear to be a trivial task. Therefore, the monohydrate adduct of α-D-glucopyranose was extracted from the X-ray crystal structure[55] and the energy evaluated together with an identical structure with the water molecule translated 50 Å away. A B3LYP/AUG-cc-pVDZ test calculation with the completely separated entities showed no significant change; that is, 50 Å separation is sufficient to model infinite separation. As shown in Figure 1F, the presence of the water molecule establishes a three-centered hydrogen-bonding scheme.

When comparing the force fields, this test yielded probably the most surprising result, as the energies calculated for this interaction proved to range from 0.85 and 1.04 kcal/mol (MM3*($\varepsilon = 80$)) and AMBER-GLYCAM 4.1, respectively) to 13.7 kcal/mol (SPACIBA). Based on the results in Table 3, our best estimate of the interaction energy in gas phase is calculated as (with the AUG-cc-pVDZ basis set):

$$\begin{aligned}\Delta E^{\text{estimate}} = {} & \Delta E^{\text{CP}}(\text{B3LYP; full system}) \\ & + \frac{1}{2}(\Delta E(\text{CCSD}(T);\text{model system 2}) \\ & + \Delta E^{\text{CP}}_{\text{GW}}(\text{CCSD}(T);\text{model system 2})) \\ & - \Delta E^{\text{CP}}(\text{B3LYP; model system 2})\end{aligned}$$

The first term is simply the energy of the full system at B3LYP level. For the B3LYP method, the CP correction works well for the water dimer (see Table 1). We therefore apply the CP-corrected results directly for this method. The second and third lines gives an estimate of the energy of the small model system at CCSD(T) level corrected for BSSE. Note that for the CCSD(T) results, the CP-corrected result is not used directly, but instead the average of the uncorrected and the CP-corrected energy, following the approach of Halkier et al.[56] For the water dimer, this correction gives a result at CCSD(T)/AUG-cc-pVDZ level, within 0.2 kcal/mol of the estimated complete basis set result. The fourth line gives the energy of the small model system at B3LYP level. Subtracting this from the CCSD(T) energy for the small model system gives an estimate of the difference between the energy at the two levels of theory. Assuming that this difference between the two methods is the same for the full system, the correction is added to the B3LYP energy for the full system, in an attempt to correct the energy calculated at B3LYP level up to CCSD(T) level for the full system. This gives $\Delta E^{\text{estimate}} = 4.9$ kcal/mol. The error is most likely smaller than the estimated 1 kcal/mol for the pure B3LYP/AUG-cc-pVDZ method. In Figure 3, this result is compared to the force field results. The QM result is roughly equal to the mean of the force field results, but considerable deviations occur.

3.4. Puckering parameters and energy difference between the north and south conformer of the methyl 5-deoxy-β-D-xylofuranoside

To investigate puckering magnitude and preference for furanose rings two conformations corresponding to opposite pucker on the pseudorotational wheel, that is, the envelope shape ^3E for the *north* pseudorotamer (Fig. 1B) and $_3E$ for the *south* pseudorotamer (Fig. 1A), were geometry optimized and the relative energy calculated. The results are listed in Table 4. The best estimate of the relative energy is that the *south* conformer is more stable than the *north* conformer by 0.66 kcal/mol (optimized at

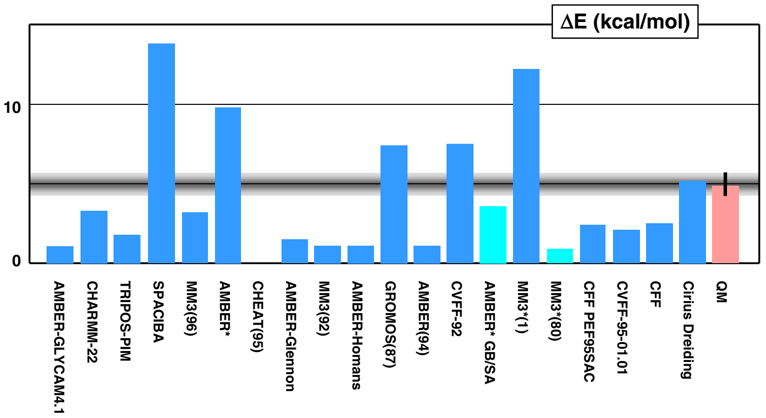

Figure 3. Interaction energy (kcal/mol) of the α-D-glucopyranose·H_2O heterodimer. Comparison of the results from 20 force field calculations with the quantum mechanical (QM) estimate using a composite method, see the text. An estimated error bar is shown for the QM result. No estimate is given for CHEAT(95) for this energy. Amber GB/SA and MM3*(80) have different shading, because they represent force fields constructed for modeling in solution, and should not be directly compared to the QM data. They are included because they were part of the original force field comparison.[7]

the B3LYP/6-311+G(d,p) level). To our knowledge, this relative energy has not been previously calculated by quantum mechanical methods, though Gordon et al.[28] have carried out a comprehensive study of conformers of methyl β-D-arabinofuranoside. Applying the smaller 6-31G(d) basis set in the geometry optimization only increased the preference for the *south* conformer to 0.79 kcal/mol. Perhaps rather characteristic of this delicate balance between axial hydrophobic groups and ring strain, very scattered values were obtained by molecular mechanics calculations ranging from 4.0 kcal/mol preference for the *north* conformer by CHARM22 to 1.4 kcal/mol preference for the *south* one by Tripos-PIM. On the experimental side, we have only vague evidence that the Gibbs free energy difference is small, but with a preference to the *north* conformational family.[7] However, such results cannot be compared to our results without ensemble averaging and condensed phase calculations.

When examining the structures in terms of the Cremer–Pople pucker parameters q_2 and ϕ_2 it is evident from Figure 4 that the pucker magnitude and angle of the furanose rings are close to most of the results from the molecular mechanics force fields in the major energy wells of the MM3 force field,[62] which is used for the background energy contours. The size of the stars corresponds roughly to the difference between the q_2 and ϕ_2 values found using the two different basis sets in the quantum mechanical calculations and is a rough estimate of an error bar. The only remarkable features are that the pucker magnitude of the *north* conformer is slightly lower than predicted by any of the MM force fields and that the pucker angle of the *south* conformer is slightly larger than predicted by most of the MM force fields. However, the contour map produced by MM3 calculations indicates that the potential energy surface is relatively flat in this area, and thus these differences may not be serious indications of force field deficiencies. The MM3 contour map was not intended to reproduce gas phase data, but it was included in the figure to give a rough overview of the potential energy surface.

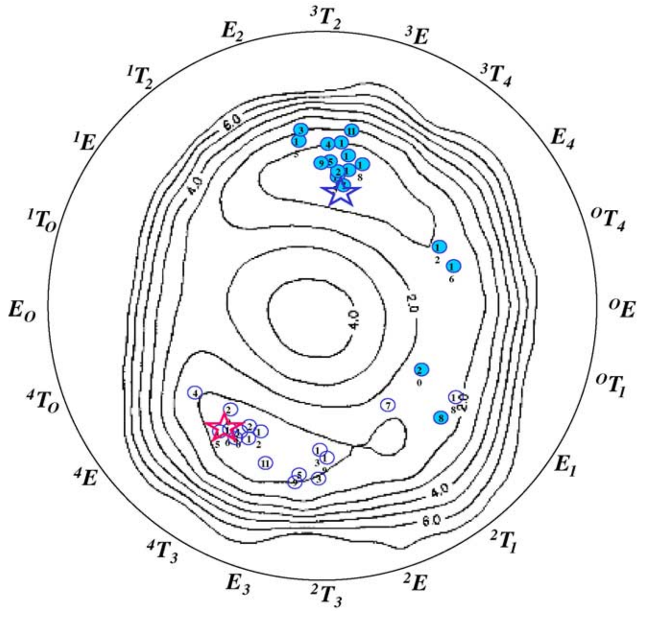

Figure 4. Dispersion in Cremer–Pople parameter space for the *north* and *south* conformer of methyl 5-deoxy-β-D-xylofuranoside. The B3LYP/AUG-cc-pVDZ//B3LYP/6-311+G(d,p) results are presented as stars, and the force field results as encircled numbers corresponding to the numbering in Perez et al.[7] Filled circles correspond to geometry optimizations started near the *north* conformation, open circles correspond to geometry optimizations started near the *south* conformation. The background contours displayed in 1 kcal/mol levels are calculated using MM3.

3.5. Structures and relative energy of methyl α- and β-D-glucopyranoside

The calculation of the geometry, conformation and energy as a function of the anomeric configuration of a glucopyranoside is a key question of interest to any carbohydrate chemist. In order to study the anomeric preference and structure, the two anomers of methyl-D-glucopyranoside in the $^4C_1$ ring conformation were geometry optimized and the relative energy calculated (see Table 5 and Fig. 5). We found that the α anomer is more stable than the β anomer by 0.82 kcal/mol. With the reiterated precautions of ensemble-averaged condensed phase modeling, this result appears to be in good accordance with the experimental anomeric ratio of methyl glucopyranoside, resulting in a preference of the α anomer by 0.42 kcal/mol.[63] Our result also agrees well with the estimated highly correlated energy difference of 1.0 kcal/mol found by Barrows et al. for glucopyranose. The different methods and the presence of the methyl group in this work may account for the difference of

Table 5. Selected structural parameters and relative energy of the *gt* conformer of methyl α-D-glucopyranoside and methyl β-D-glucopyranoside determined at B3LYP/AUG-cc-pVDZ//B3LYP/6-31G(d) level

	Methyl α-D-glucopyranoside	Methyl β-D-glucopyranoside
Bond lengths (Å)		
O-5–C-1	1.41	1.42
C-1–O-1	1.41	1.39
Angles (deg)		
O-5–C-1–O-1	113.4	109.8
C-1–O-1–CMe	114.2	114.6
Dihedral angles (deg)		
C-5–O-5–C-1–O-1	62.8	179.2
C-5–O-5–C-1–H-1	−177.2	57.9
O-5–C-1–O-1–CMe	65.4	−70.9
H-1–C-1–O-1–CMe	−51.6	49.7
Energy		
$\Delta E_{\alpha-\beta}$ (kcal/mol)	−0.82	

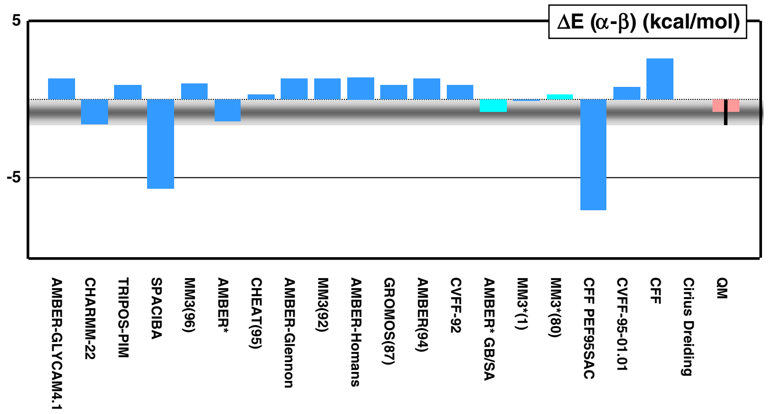

Figure 5. Energy difference (kcal/mol) between methyl α-D-glucopyranoside and methyl β-D-glucopyranoside. Comparison of the results from 20 force field calculations with the B3LYP/AUG-cc-pVDZ//B3LYP/6-31G(d) (QM) result. An estimated error bar is shown for the QM result. Amber GB/SA and MM3*(80) have different shading, because they represent force fields constructed for modeling in solution, and should not be directly compared to the QM data. They are included because they were part of the original force field comparison.[7]

0.2 kcal/mol, indicating that the anomeric effect might be smaller for methyl glucopyranoside than for glucopyranose. The results of Lii et al.[12] support this hypothesis. Using B3LYP/6-311++G(2d,2p)//6-31G(d,p), they found an energy difference of 1.39 kcal/mol. Ma et al.[10] have estimated this energy difference at the B3LYP/6-31G(d,p) level to be 1.7 kcal/mol. Similarly, Salzner and Schleyer[25] found energies of 1.7 and 1.4 kcal/mol at the HF/6-31G(d) level for methyl glucopyranoside and glucopyranose, respectively. The latter seems to contradict the notion that the anomeric effect is smaller for methyl glucopyranoside than for glucopyranose, but this is probably because relatively small basis sets were used in those studies. A thorough study of the anomeric and exo-anomeric effect was performed by Tvaroska and Carver[14–17,19] on a smaller model system, 2-methoxy-tetrahydropyran, giving an energy difference of 0.65 kcal/mol, in very good agreement with our result.

In contrast to the quantum mechanical methods, the molecular mechanics force fields fail to predict this anomeric effect, since most of them yielded a small energy preference for the β-anomer (see Fig. 5). Only CHARM-22, AMBER for MacroModel and AMBER* together with GB/SA solvation model yield an energy difference close to the value found in this study. The good accordance of the latter force field is explained by the fact that a similar procedure was used to optimize the force field parameter terms.[64]

With respect to the structural aspects of the anomeric effect, our calculations gave a C-1–O-1 bond lengths of α: 1.41 Å/β: 1.39 Å and valence angles O-5–C-1–O-1 of α: 113.4°/β: 109.8°, in agreement with the experimentally shorter C-1–O-1 bonds and smaller O-5–C-1–O-1 angles observed in the solid state for the β-anomers of aldopyranosides.[65,66] For both the α- and β-anomer, most of the force fields give longer C-1–O-1 bond lengths, though all are within 0.05 Å of the QM result. Interestingly, only 7 of the 20 force fields reproduce the decrease in C-1–O-1 bond length from the α to the β anomer. For the O-5–C-1–O1 and C-1–O-1–CMe valence angles all force fields give results close to the QM result (within 7°). All except two (SPACIBA and CVFF-92) reproduce the decrease in valence angle from the α to the β-anomer. The calculated glycosidic valence angles O-5–C-1–CMe were α: 114.2°/β: 114.6° rather invariant to the anomer conformation and in agreement with most force field calculations and experimental X-ray evidences. The dihedral angles defining the anomer (C-5–O-5–C-1–O-1 and C-5–O-5–C-1–H-1) are also very similar for all force fields and within 6° of the QM result. The dihedral angles giving the orientation of the methyl group at the ‘glycosidic linkage’ (O-5–C-1–O-1–CMe and H-1–C-1–O-1–CMe), however, display larger variations among the force fields. Most give values within 10° of the QM result, and the largest deviations are 16–19° (found with CHARMM-22 and SPACIBA). In conclusion, the calculations indicate that most molecular mechanics force fields give acceptable geometries around the anomeric center of the glucopyranosides, whereas they generally fail to predict the correct sign for the energy difference between the two anomers.

3.6. Structures and relative energy of primary hydroxyl rotamers of methyl α-D-glucopyranoside and methyl α-D-galactopyranoside

The population ratio of the rotational conformers around the C-5–C-6 torsion is an important issue, as it may govern inter-molecular interactions. The three staggered conformations of the hydroxymethyl group for the two epimers in the $^4C_1$ ring conformation were subjected to geometry optimization and calculation of relative energies between rotamers at the B3LYP/AUG-cc-pVDZ//B3LYP/6-31G(d) level. These three

Table 6. O-5–C-5–C-6–O-6 dihedral angle and relative energies of hydroxymethyl group rotamers in methyl α-D-glucopyranoside and methyl α-D-galactopyranoside determined at B3LYP/AUG-cc-pVDZ//B3LYP/6-31G(d) level

	Methyl α-D-glucopyranoside			Methyl α-D-galactopyranoside		
	gg	*gt*	*tg*	*gg*	*gt*	*tg*
Relative energy (kcal/mol)	0.13	0.00	0.15	0.64	0.00	0.77
Dihedral angle (deg) O-5–C-5–C-6–O-6	−57.0	58.0	168.3	−57.4	58.3	−164.5

orientations are referred as to gauche–gauche (*gg*), gauche–trans (*gt*), and trans–gauche (*tg*), with values of −60°, +60°, and 180° for the O-5–C-5–C-6–O-6 torsion angle, respectively. The results are listed in Table 6. For methyl α-D-glucopyranoside the three energy minima are very close in energy and in the order *gg*, *gt*, *tg* (0.13, 0.00, 0.15 kcal/mol), indicating the existence of a small gauche effect even in the gas phase, though this is within the accuracy of the calculations. For glucopyranose, Brown and Wladkowski[9] found a different order and larger energy differences (0.02, 0.56, 0.00 kcal/mol) using the MP2/6-31G(d) method, which they considered their best result. Ma et al.[10] similarly found the order (0.55, 0.81, 0.00 kcal/mol) using the B3LYP/6-31G(d,p) method. The difference from the results found in this work is most likely due to the inclusion of diffuse functions in the basis sets used here. This is supported by Lii et al.[12] and Barrows et al.[11] who found results similar to ours (0.10, 0.00, 0.12 kcal/mol and 0.0, 0.1, 0.2 kcal/mol, respectively) using B3LYP/6-311++ G(2d,2p)//B3LYP/6-31G(d,p) and a composite energy method, respectively. Surprisingly, Tvaroska et al.[37] found somewhat larger energy differences between the three rotamers, at a similar level of theory B3LYP/6-311++G(d,p)//B3LYP/6-31G(d) using the Jaguar program.[67] We have performed a series of test calculations with the method used by Tvaroska et al. with Gaussian, as well as with Jaguar, and obtained the results in Table 6 in all cases. Thus, the most likely explanation for the discrepancies is that the molecular structures in Tvaroska et al.[37] differ from the structures in this work and probably also from the structures in the work by Lii et al.[12] and Barrows et al.[11]

From NMR experiments, the rotamer population has been estimated from pro*R* and pro*S* couplings from the C-6 protons to present a *gg*:*gt*:*tg* ratio of 57:38:5 for methyl α-D-glucopyranoside.[68] This indicates that the relative free energies of the three rotamers are (0.00, 0.24, 1.44) kcal/mol under the experimental conditions. The discrepancy between quantum mechanical calculations and the NMR data can be attributed to inaccuracies of the calculations, lack of ensemble averaging or subtle condensed phase effects. It is well known that molecular mechanics force fields applied in vacuum do not accurately model this population distribution.[69] In this context of gas-phase calculations, it is more relevant to evaluate and compare the magnitude of the energy difference between the three staggered hydroxymethyl conformers with the quantum mechanical results. In this case (Fig. 6), the results indicate that most MM force fields exaggerate the energy difference when compared to the QM result.

For the calculations on methyl α-D-galactopyranoside, we found the same ordering but with a slightly larger energy differentiation: *gg*, *gt*, *tg* with relative energies of 0.64, 0.00, 0.77 kcal/mol. This is in good agreement with the results of Lii et al.[12] for galactopyranose. The additional methyl group in our systems is

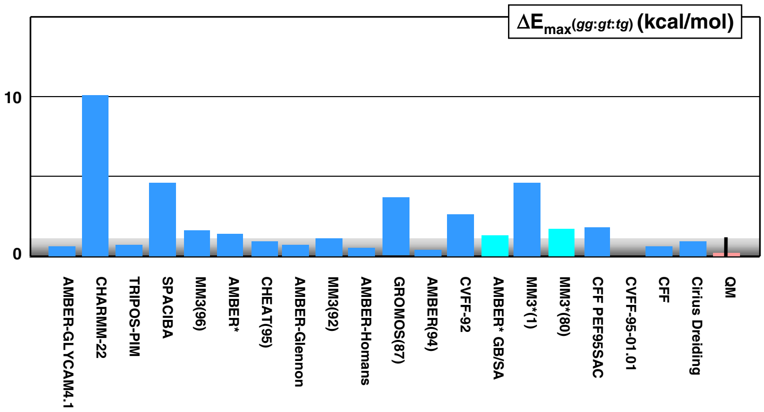

Figure 6. Maximum energy difference between hydroxymethyl rotamers of methyl α-D-glucopyranoside. Comparison of the results from 20 force field calculations with the B3LYP/AUG-cc-PVDZ//B3LYP/6-31G(d) (QM) result. An estimated error bar is shown for the QM result. Amber GB/SA and MM3*(80) have different shading, because they represent force fields constructed for modeling in solution, and should not be directly compared to the QM data. They are included because they were part of the original force field comparison.[7]

not expected to play a significant role for the relative energies of the three conformers. From NMR experiments, the rotamer population was estimated to present a *gg*:*gt*:*tg* ratio of 14:47:39,[68] indicating the relative free energies of the three rotamers are (0.72, 0.00, 0.11) kcal/mol under the experimental conditions. The conformational behavior may be dependent on delicately balanced and unique solvent interactions.[2] Most of the force fields also predict relatively small energy differences between the three conformations. However, some give larger energy differences, most notably 4.5 and 7.9 kcal/mol from GROMOS(87) and MM3*($\varepsilon = 1$), respectively.

The force fields display considerable variations in the O-5–C-5–C-6–O-6 dihedral angle for each of the three rotamers of both methyl α-D-glucopyranoside and methyl α-D-galactopyranoside. The largest deviation from the QM result is 30°, found for the *tg* rotamer of α-D-galactopyranoside optimized with CHARMM-22, but SPACIBA, AMBER-Glennon, AMBER-Homans, MM3(92), and MM3*($\varepsilon = 1$) also give deviations of more than 20°. These large variations might reflect a relatively flat potential energy surface around the rotamer minima, that is, a real physical effect. However, the energy barriers of 4–7 kcal/mol between the rotamers found by Brown and Wladkowski[9] indicate that this is not the case. It is more likely that the reason is different approaches to parameterization of dihedral terms in the force fields and thus deficiencies in some of the force fields.

3.7. The Sammon map

In the original force field benchmark test, we constructed a PCA landscape to which the force fields could be related on a relative scale. The result was that a clear clustering of the performances of the Amber-based force fields was observed and to some extent a clustering of MM3-based force fields albeit the performance of the partial charge-based MacroModel MM3* was significantly different from the bond dipole electrostatics in the original MM3.

In the present study, we have taken this relative analysis one step further by constructing a Sammon map from the PCA scores, allowing the entire systematic variation in the force field performances to be visualized in a simple 2D plot. This analysis is based on the assumption that the performance of the 20 different carbohydrate force fields can be evaluated on the basis of 32 molecular quantities of the monosaccharides in the test set (7 energies, 4 bond lengths, 4 valence angles, 13 dihedral angles and 4 pucker parameters) contributing with equal weight.

The Sammon map of the force field performances is displayed in Figure 7. By chance, the quantum mechanical (QM) result is almost at the center of the map, due to the fact that the result of the quantum mechanical calculations in many test cases is near the mean of the force field predictions. As in the previously published PCA map,[7] we found a cluster of Amber-based force field to the *north* and a trend of MM3-based force fields to the *south-east*. The force field that is closest in performance to the quantum mechanical result is the extended atom CHEAT force field. Ironically, the original Amber force field is the closest of the Amber family force fields to the quantum mechanical calculation. From the map, it is also observed that the MM3* parameterization performs better when the influence of the electrostatic potential is dampened by a large dielectric constant. The QM results are, of course, not exact. It was investigated how variations of the QM

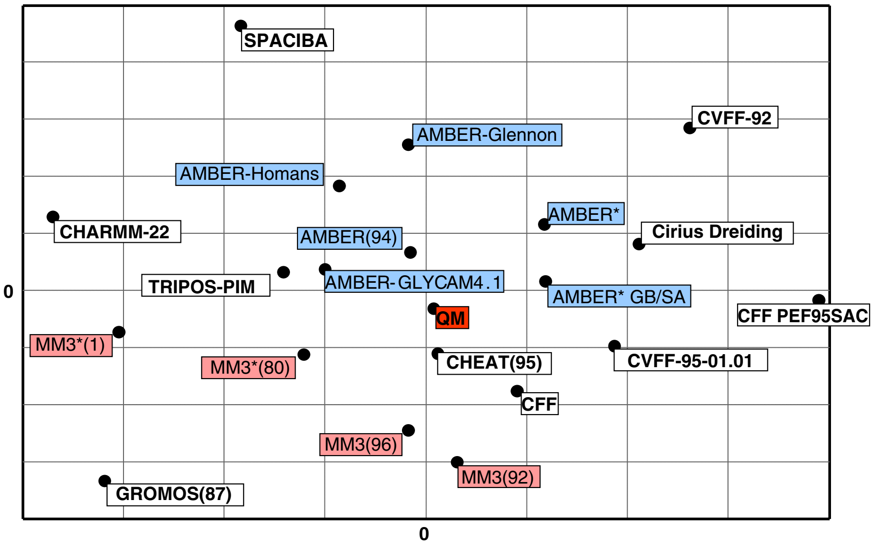

Figure 7. Sammon map: An overall estimate of the performance of the molecular mechanics force fields. Proximity to the quantum mechanical (QM) position indicates best overall performance.

results within estimated uncertainties (0.5 kcal/mol, 0.1 Å, 5° and 10° for energies, bond lengths, valence angles and dihedral angles) affected the plot. The QM position changed considerably in the y-direction, but almost not at all in the x-direction. Thus, a large distance between the QM position and the respective force field positions in the x-direction indicates serious deficiencies of the force field, whereas differences along the y-direction are not equally significant. The conclusion is, therefore, that CHEAT(95), Amber(94), Amber-Glennon, MM3(96), MM3(92) and to some extent CFF have the best overall performances. It should be kept in mind, however, that this does not necessarily mean that these force fields will also perform well for larger systems.

4. Conclusions

With this study, we have confirmed three major problems with the second-generation carbohydrate force fields. First of all, the interaction between α-D-glucopyranose and its crystalline hydrate is estimated to be 4.9 kcal/mol, and most force fields are strongly in error in this respect. Secondly, the quantum mechanical calculations show that the staggered populations around the C-5–C-6 linkage in methyl α-D-glucopyranoside and methyl α-D-galactopyranoside are separated by small energy differences in the gas phase, a delicate balance that may be altered upon solvation. Most force fields give too large energy differences between the *gg*:*gt*:*tg* rotamers. Most force fields fail to predict the correct anomer ranking for a classical carbohydrate example: methyl α/β-D-glucopyranoside. In terms of structure the force fields generally perform well, with bond lengths within 0.05 Å, angles within 7° and most dihedral angles within 10° of the QM results. However, larger deviations up to 30° occur for the O-5–C-5–C-6–O-6 dihedral angle of the three rotamers of methyl α-D-glucopyranoside and methyl α-D-galactopyranoside.

It is perhaps a rather disappointing result that three of the most crucial and distinct features of carbohydrates are erroneously predicted by the majority of the second generation molecular mechanics force fields. However, it should be noted that, for example, the α/β anomer energy preference is not a primary target in some force field developments, as they cannot mutate in classical molecular dynamics calculations. As an example, the Amber GLYCAM force field, which is developed with extensive use of ab initio calculations uses a charge set-up that is optimized for use with explicit waters and uses different charges for the α- and β-anomers. The result is that, while Glycam reproduces ΔG_{SOLV} in explicit waters,[70] it overestimates the ΔE energy difference in the gas phase.[71] Nevertheless, it would be regarded as a sound property of the force fields if the three above-mentioned features of carbohydrates were modeled in better accordance with the quantum mechanical results.

In conclusion, this study indicates that there is ample room for development of new third-generation carbohydrate force fields. It will be interesting to investigate whether, for example, a more appropriate modeling of the charge density (including lone pairs and perhaps also multipole charges and polarizabilities) will remedy the problems and provide adequate models of the fine stereoelectronic and energetic details of carbohydrates.

Acknowledgements

The Centre for Advanced Food Studies is acknowledged for supporting molecular modeling and The Carlsberg Foundation is acknowledged for financial support. We thank Per-Ola Norrby for performing test calculations with the Jaguar program and Gilda Kischinovsky for assistance with the manuscript. We also thank Muhammad Sharaz Butt for participating in the project.

References

1. Tvaroska, I.; Bleha, T. *Adv. Carbohydr. Chem. Biochem.* **1989**, *47*, 45–123.
2. Engelsen, S. B.; Monteiro, C.; Hervé du Penhoat, C.; Pérez, S. *Biophys. Chem.* **2001**, *93*, 103–127.
3. Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127–8134.
4. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
5. Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230–252.
6. van Gunsteren, W. F.; Berendsen, H. J. C. Groningen Molecular Simulation (GROMOS) Program Package, 1987. Groningen, The Netherlands, Nijenborgh 16, University of Groningen.
7. Pérez, S.; Imberty, A.; Engelsen, S. B.; Gruza, J.; Mazeau, K.; Jiménez-Barbero, J.; Poveda, A.; Espinoza, J. F.; van Eijck, B. P.; Johnson, G.; French, A. D.; Kouwijzer, M. L. C. E.; Grootenhuis, P. D. J.; Bernardi, A.; Raimondi, L.; Senderowitz, H.; Durier, V.; Vergoten, G.; Rasmussen, K. *Carbohydr. Res.* **1998**, *314*, 141–155.
8. Momany, F. A.; Willett, J. L. *J. Comput. Chem.* **2000**, *21*, 1204–1219.
9. Brown, J. W.; Wladkowski, B. D. *J. Am. Chem. Soc.* **1996**, *118*, 1190–1193.
10. Ma, B. Y.; Schaefer, H. F.; Allinger, N. L. *J. Am. Chem. Soc.* **1998**, *120*, 3411–3422.
11. Barrows, S. E.; Storer, J. W.; Cramer, C. J.; French, A. D.; Truhlar, D. G. *J. Comput. Chem.* **1998**, *19*, 1111–1119.
12. Lii, J.-H.; Ma, B. Y.; Allinger, N. L. *J. Comput. Chem.* **1999**, *20*, 1593–1603.
13. Chung-Phillips, A.; Chen, Y. Y. *J. Phys. Chem. A* **1999**, *103*, 953–964.
14. Tvaroska, I.; Carver, J. P. *J. Phys. Chem.* **1994**, *98*, 6452–6458.
15. Tvaroska, I.; Carver, J. P. *J. Phys. Chem.* **1994**, *98*, 9477–9485.

16. Tvaroska, I.; Carver, J. P. *J. Phys. Chem.* **1996**, *100*, 11305–11313.
17. Tvaroska, I.; Carver, J. P. *J. Phys. Chem. B* **1997**, *101*, 2992–2999.
18. Tvaroska, I.; Carver, J. P. *J. Mol. Struct. (Theochem)* **1997**, *395*, 1–13.
19. Tvaroska, I.; Carver, J. P. *Carbohydr. Res.* **1998**, *309*, 1–9.
20. Csonka, G. I.; Elias, K.; Csizmadia, I. G. *J. Comput. Chem.* **1997**, *18*, 330–342.
21. Csonka, G. I.; Elias, K.; Kolossvary, I.; Sosa, C. P.; Csizmadia, I. G. *J. Phys. Chem. A* **1998**, *102*, 1219–1229.
22. Barrows, S. E.; Dulles, F. J.; Cramer, C. J.; French, A. D.; Truhlar, D. G. *Carbohydr. Res.* **1995**, *276*, 219–251.
23. Wiberg, K. B.; Marquez, M. *J. Am. Chem. Soc.* **1994**, *116*, 2197–2198.
24. Jorgensen, W. L.; Morales de Tirado, P. I.; Severance, D. L. *J. Am. Chem. Soc.* **1994**, *116*, 2199–2200.
25. Salzner, U.; Schleyer, P. v. R. *J. Org. Chem.* **1994**, *59*, 2138–2155.
26. Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1993**, *115*, 5745–5753.
27. Wiberg, K. B.; Murcko, M. A. *J. Am. Chem. Soc.* **1989**, *111*, 4821–4828.
28. Gordon, M. T.; Lowary, T. L.; Hadad, C. M. *J. Org. Chem.* **2000**, *65*, 4954–4963.
29. Mulroney, B.; Peel, J. B.; Traeger, J. C. *J. Mass Spectrom.* **1999**, *34*, 544–553.
30. French, A. D.; Kelterer, A. M.; Johnson, G. P.; Dowd, M. K.; Cramer, C. J. *J. Comput. Chem.* **2001**, *22*, 65–78.
31. Cramer, C. J.; Kelterer, A. M.; French, A. D. *J. Comput. Chem.* **2001**, *22*, 1194–1204.
32. Gregurick, S. K.; Kafafi, S. A. *Carbohydr. Chem.* **1999**, *18*, 867–890.
33. Molteni, C.; Parrinello, M. *Chem. Phys. Lett.* **1997**, *275*, 409–413.
34. Jockusch, R. A.; Talbot, F. O.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2003**, *5*, 1502–1507.
35. Talbot, F. O.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2002**, *4*, 3562–3565.
36. Klein, R. A. *J. Am. Chem. Soc.* **2002**, *124*, 13931–13937.
37. Tvaroska, I.; Taravel, F. R.; Utille, J. P.; Carver, J. P. *Carbohydr. Res.* **2002**, *337*, 353–367.
38. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Menucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Ragavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. GAUSSIAN 98. 1998. Pittsburg PA, Gaussian.
39. Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724.
40. Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.
41. Hariharan, P. C.; Pople, J. A. *Theoretica Chim. Acta* **1973**, *28*, 213–222.
42. Hariharan, P. C.; Pople, J. A. *Mol. Phys.* **1974**, *27*, 209–214.
43. Gordon, M. S. *Chem. Phys. Lett.* **1980**, *76*, 163–168.
44. McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639–5648.
45. Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650–654.
46. Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. *J. Comput. Chem.* **1983**, *4*, 294–301.
47. Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
48. Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
49. Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1994**, *98*, 1358–1371.
50. Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622.
51. Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479–483.
52. Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev.* **1988**, *37*, 785–789.
53. Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372–1377.
54. Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
55. Hough, E.; Neidle, H.; Rogers, D.; Troughton, P. G. H. *Acta Crystallogr. Sect. B, Struct. Sci.* **1973**, *29*, 365–367.
56. Halkier, A.; Koch, H.; Jørgensen, P.; Christiansen, O.; Nielsen, I. M. B.; Helgaker, T. *Theor. Chem. Acc.* **1997**, *97*, 150–157.
57. Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
58. Hotelling, H. *J. Educ. Psych.* **1933**, *24*, 417–441.
59. Sammon, J. W. *IEEE Trans. Comput.* **1969**, *18*, 401–409.
60. Csonka, G. I. *J. Mol. Struct. (Theochem)* **2002**, *584*, 1–4.
61. Cremer, D.; Pople, J. A. *J. Am. Chem. Soc.* **1975**, *97*, 1354–1358.
62. Allinger, N. L.; Rahman, M.; Lii, J.-H. *J. Am. Chem. Soc.* **1990**, *112*, 8293–8307.
63. Smirnyagin, V.; Bishop, C. T. *Can. J. Chem.* **1968**, *46*, 3085–3090.
64. Senderowitz, H.; Parish, C.; Still, W. C. *J. Am. Chem. Soc.* **1996**, *118*, 2078–2086.
65. Jeffrey, G. A.; Takagi, S. *Acta Crystallogr. Sect. B Struct. Sci.* **1977**, *33*, 738–742.
66. Takagi, S.; Jeffrey, G. A. *Acta Cryst.* **1979**, *B35*, 902–906.
67. Schrödinger. Jaguar [v. 5.0]. 2003. Portland, OR, Schrödinger. Ref Type: Computer Program.
68. Nishida, Y.; Ohrui, H.; Meguro, H. *Tetrahedron Lett.* **1984**, *25*, 1575–1578.
69. Engelsen, S. B.; Pérez, S.; Braccini, I.; Hervé du Penhoat, C. *J. Comput. Chem.* **1995**, *16*, 1096–1119.
70. Woods, R. J. *Glycoconjugate J.* **1998**, *15*, 209–216.
71. Basma, M.; Sundara, S.; Calgan, D.; Vernali, T.; Woods, R. J. *J. Comput. Chem.* **2001**, *22*, 1125–1137.